



OIIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/017,372A

DATE: 05/02/2002

TIME: 13:00:06

Input Set : N:\Crf3\04152002\J017372.raw
 Output Set: N:\CRF3\05022002\J017372A.raw

**Does Not Comply
 Corrected Diskette Needed**

C--> 1 <110> APPLICANT: Wolfram, Lawrence
 2 Letterio, John
 3 <120> TITLE OF INVENTION: FUNCTIONALIZED TGF-BETA FUSION PROTEINS
 4 <130> FILE REFERENCE: 4239-61302
 5 <140> CURRENT APPLICATION NUMBER: US/10/017,372A
 6 <141> CURRENT FILING DATE: 2000-10-19
 7 <150> PRIOR APPLICATION NUMBER: 60/242,292
 8 <151> PRIOR FILING DATE: 2000-10-20
 9 <160> NUMBER OF SEQ ID NOS: 39
 10 <170> SOFTWARE: PatentIn version 3.1
 12 <210> SEQ ID NO: 1
 13 <211> LENGTH: 28
 14 <212> TYPE: DNA
 15 <213> ORGANISM: synthetic oligonucleotide
 16 <400> SEQUENCE: 1
 17 ggagagatct ggtaccgaga tggcgcct
 19 <210> SEQ ID NO: 2
 20 <211> LENGTH: 42
 21 <212> TYPE: DNA
 22 <213> ORGANISM: synthetic oligonucleotide
 23 <400> SEQUENCE: 2
 24 ataagaattg cggccgcttt aatcgatccc aagtgggctt gg
 26 <210> SEQ ID NO: 3
 27 <211> LENGTH: 48
 28 <212> TYPE: DNA
 29 <213> ORGANISM: synthetic oligonucleotide
 30 <400> SEQUENCE: 3
 31 gactacaagg atgacgacga caaggccctg gataccaact actgcttc
 33 <210> SEQ ID NO: 4
 34 <211> LENGTH: 45
 35 <212> TYPE: DNA
 36 <213> ORGANISM: synthetic oligonucleotide
 37 <400> SEQUENCE: 4
 38 cttgtcgtcg tcatccttgt agtctcggcg gtgccgggag ctgtg
 40 <210> SEQ ID NO: 5
 41 <211> LENGTH: 45
 42 <212> TYPE: DNA
 43 <213> ORGANISM: synthetic oligonucleotide
 44 <400> SEQUENCE: 5
 45 gactacaagg atgacgacga caggagaaga actgctgcgt gcggc
 47 <210> SEQ ID NO: 6
 48 <211> LENGTH: 45
 49 <212> TYPE: DNA

See pp 1-2, 4-5

See item 10 on Error Summary sheet

28

42

48

45

45

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/017,372A

DATE: 05/02/2002
 TIME: 13:00:06

Input Set : N:\Crif3\04152002\J017372.raw
 Output Set: N:\CRF3\05022002\J017372A.raw

50 <213> ORGANISM: synthetic oligonucleotide
 51 <400> SEQUENCE: 6 45
 52 cttgtcgtcg tcataccttgt agtctcggcg gtgccgggag ctgtg
 54 <210> SEQ ID NO: 7
 55 <211> LENGTH: 8
 56 <212> TYPE: PRT
 57 <213> ORGANISM: synthetic oligopeptide
 58 <400> SEQUENCE: 7
 59 Asp Tyr Lys Asp Asp Asp Asp Lys
 60 1 5
 62 <210> SEQ ID NO: 8
 63 <211> LENGTH: 1197
 64 <212> TYPE: DNA
 65 <213> ORGANISM: fusion
 66 <220> FEATURE:
 67 <221> NAME/KEY: CDS
 68 <222> LOCATION: (1)..(1197)
 69 <223> OTHER INFORMATION:
 70 <220> FEATURE:
 71 <221> NAME/KEY: misc_feature
 72 <222> LOCATION: (278)..(279)
 73 <223> OTHER INFORMATION: Maturation cleavage site
 74 <400> SEQUENCE: 8 48
 75 atg gcg cct tcg ggg ctg cgg ctc ttg ccg ctg ctg ctg ccg ctg ctg
 76 Met Ala Pro Ser Gly Leu Arg Leu Leu Pro Leu Leu Leu Pro Leu Leu
 77 1 5 10 15 96
 78 tgg ctg cta gtg ctg acg cct ggc cgg ccg gcc gcc gga ctg tcc acc
 79 Trp Leu Leu Val Leu Thr Pro Gly Arg Pro Ala Ala Gly Leu Ser Thr
 80 20 25 30
 81 tgc aag acc atc gac atg gag ctg gtg aag cgg aag cgc atc gag gcc 144
 82 Cys Lys Thr Ile Asp Met Glu Leu Val Lys Arg Lys Arg Ile Glu Ala
 83 35 40 45
 84 att cgc ggc cag att ctg tcc aag ctt cgg ctt gcc agc ccc ccg agc 192
 85 Ile Arg Gly Gln Ile Leu Ser Lys Leu Arg Leu Ala Ser Pro Pro Ser
 86 50 55 60
 87 cag ggg gac gtg ccg ccc ggc ccg ctg cct gag gca gta ctg gct ctt 240
 88 Gln Gly Asp Val Pro Pro Gly Pro Leu Pro Glu Ala Val Leu Ala Leu
 89 65 70 75 80
 90 tac aac agt acc cgc gac cgg gta gcc ggg gaa agt gtc gaa ccg gag 288
 91 Tyr Asn Ser Thr Arg Asp Arg Val Ala Gly Glu Ser Val Glu Pro Glu
 92 85 90 95
 93 ccc gag cca gag gcg gac tac tac gcc aag gag gtc acc cgc gtg cta 336
 94 Pro Glu Pro Glu Ala Asp Tyr Tyr Ala Lys Glu Val Thr Arg Val Leu
 95 100 105 110
 96 atg gtg gaa agc ggc aac caa atc tat gat aaa ttc aag ggc acc ccc 384
 97 Met Val Glu Ser Gly Asn Gln Ile Tyr Asp Lys Phe Lys Gly Thr Pro
 98 115 120 125
 99 cac agc tta tat atg ctg ttc aac acg tcg gag ctc cgg gaa gcg gtg 432
 100 His Ser Leu Tyr Met Leu Phe Asn Thr Ser Glu Leu Arg Glu Ala Val

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/017,372A

DATE: 05/02/2002
 TIME: 13:00:06

Input Set : N:\Crf3\04152002\J017372.raw
 Output Set: N:\CRF3\05022002\J017372A.raw

101	130	135	140		480
102	ccg gaa cct gta ttg ctc tct cgg gca gag ctg cgc ctg ctg agg ctc				
103	Pro Glu Pro Val Leu Leu Ser Arg Ala Glu Leu Arg Leu Leu Arg Leu				
104	145	150	155	160	528
105	aag tta aaa gtg gag cag cac gtg gag cta tac cag aaa tac agc aat				
106	Lys Leu Lys Val Glu Gln His Val Glu Leu Tyr Gln Lys Tyr Ser Asn				
107	165	170		175	576
108	gat tcc tgg cgc tac ctc agc aac cgg ctg ctg gcc ccc agt gac tca				
109	Asp Ser Trp Arg Tyr Leu Ser Asn Arg Leu Leu Ala Pro Ser Asp Ser				
110	180	185	190		624
111	ccg gag tgg ctg tcc ttt gat gtc acc gga gtt gtg cgg cag tgg ctg				
112	Pro Glu Trp Leu Ser Phe Asp Val Thr Gly Val Val Arg Gln Trp Leu				
113	195	200	205		672
114	acc cgc aga gag gct ata gag ggt ttt cgc ctc agt gcc cac tct tcc				
115	Thr Arg Arg Glu Ala Ile Glu Gly Phe Arg Leu Ser Ala His Ser Ser				
116	210	215	220		720
117	tct gac agc aaa gat aac aca ctc cac gtg gaa att aac ggg ttc aat				
118	Ser Asp Ser Lys Asp Asn Thr Leu His Val Glu Ile Asn Gly Phe Asn				
119	225	230	235	240	768
120	tct ggc cgc cgg ggt gac ctg gcc acc att cac ggc atg aac cgg ccc				
121	Ser Gly Arg Arg Gly Asp Leu Ala Thr Ile His Gly Met Asn Arg Pro				
122	245	250	255		816
123	ttc ctg ctc ctc atg gcc acc ccg ctg gag agg gcc cag cac ctg cac				
124	Phe Leu Leu Leu Met Ala Thr Pro Leu Glu Arg Ala Gln His Leu His				
125	260	265	270		864
126	agc tcc cgg cac cgc cga gac tac aag gat gac gac gac aag gcc ctg				
127	Ser Ser Arg His Arg Arg Asp Tyr Lys Asp Asp Asp Asp Lys Ala Leu				
128	275	280	285		912
129	gat acc aac tac tgc ttc agc tcc acg gag aag aac tgc tgc gtg cgg				
130	Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys Val Arg				
131	290	295	300		960
132	cag ctc tac att gac ttc cgg aag gac ctg ggc tgg aag tgg att cat				
133	Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp Ile His				
134	305	310	315	320	1008
135	gaa ccc aag ggc tac cat gcc aat ttc tgc ctg ggg ccc tgt ccc tac				
136	Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly Pro Cys Pro Tyr				
137	325	330	335		1056
138	atc tgg agc cta gac act cag tac agc aag gtc ctg gct ctg tac aac				
139	Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu Tyr Asn				
140	340	345	350		1104
141	cag cac aac ccg ggc gcg tgc gcg gcg ccg tgc tgc gtg ccg cag gcg				
142	Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys Val Pro Gln Ala				
143	355	360	365		1152
144	ctg gag cca ctg ccc atc gtg tac gtg ggc cgc aag ccc aag gtg				
145	Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro Lys Val				
146	370	375	380		1197
147	gag cag ctg tcc aac atg atc gtg cgt tcc tgc aag tgc agc tga				
148	Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser				
149	385	390	395		

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/017,372A

DATE: 05/02/2002
TIME: 13:00:06

Input Set : N:\Crf3\04152002\J017372.raw
Output Set: N:\CRF3\05022002\J017372A.raw

```

151 <210> SEQ ID NO: 9
152 <211> LENGTH: 398
153 <212> TYPE: PRT
154 <213> ORGANISM: fusion
155 <220> FEATURE:
156 <221> NAME/KEY: misc_feature
157 <222> LOCATION: (278)..(279)
158 <223> OTHER INFORMATION: Maturation cleavage site
159 <400> SEQUENCE: 9
160 Met Ala Pro Ser Gly Leu Arg Leu Leu Pro Leu Leu Leu Pro Leu Leu
161 1 5 10 15
162 Trp Leu Leu Val Leu Thr Pro Gly Arg Pro Ala Ala Gly Leu Ser Thr
163 20 25 30
164 Cys Lys Thr Ile Asp Met Glu Leu Val Lys Arg Lys Arg Ile Glu Ala
165 35 40 45
166 Ile Arg Gly Gln Ile Leu Ser Lys Leu Arg Leu Ala Ser Pro Pro Ser
167 50 55 60
168 Gln Gly Asp Val Pro Pro Gly Pro Leu Pro Glu Ala Val Leu Ala Leu
169 65 70 75 80
170 Tyr Asn Ser Thr Arg Asp Arg Val Ala Gly Glu Ser Val Glu Pro Glu
171 85 90 95
172 Pro Glu Pro Glu Ala Asp Tyr Tyr Ala Lys Glu Val Thr Arg Val Leu
173 100 105 110
174 Met Val Glu Ser Gly Asn Gln Ile Tyr Asp Lys Phe Lys Gly Thr Pro
175 115 120 125
176 His Ser Leu Tyr Met Leu Phe Asn Thr Ser Glu Leu Arg Glu Ala Val
177 130 135 140
178 Pro Glu Pro Val Leu Leu Ser Arg Ala Glu Leu Arg Leu Leu Arg Leu
179 145 150 155 160
180 Lys Leu Lys Val Glu Gln His Val Glu Leu Tyr Gln Lys Tyr Ser Asn
181 165 170 175
182 Asp Ser Trp Arg Tyr Leu Ser Asn Arg Leu Leu Ala Pro Ser Asp Ser
183 180 185 190
184 Pro Glu Trp Leu Ser Phe Asp Val Thr Gly Val Val Arg Gln Trp Leu
185 195 200 205
186 Thr Arg Arg Glu Ala Ile Glu Gly Phe Arg Leu Ser Ala His Ser Ser
187 210 215 220
188 Ser Asp Ser Lys Asp Asn Thr Leu His Val Glu Ile Asn Gly Phe Asn
189 225 230 235 240
190 Ser Gly Arg Arg Gly Asp Leu Ala Thr Ile His Gly Met Asn Arg Pro
191 245 250 255
192 Phe Leu Leu Leu Met Ala Thr Pro Leu Glu Arg Ala Gln His Leu His
193 260 265 270
194 Ser Ser Arg His Arg Arg Asp Tyr Lys Asp Asp Asp Asp Lys Ala Leu
195 275 280 285
196 Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys Val Arg
197 290 295 300
198 Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp Ile His
199 305 310 315 320

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/017,372A

DATE: 05/02/2002

TIME: 13:00:06

Input Set : N:\Crif3\04152002\J017372.raw

Output Set: N:\CRF3\05022002\J017372A.raw

```

200   Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly Pro Cys Pro Tyr
201               325               330               335
202   Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu Tyr Asn
203               340               345               350
204   Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys Val Pro Gln Ala
205               355               360               365
206   Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro Lys Val
207               370               375               380
208   Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser
209               385               390               395
211 <210> SEQ ID NO: 10
212 <211> LENGTH: 362
213 <212> TYPE: DNA
214 <213> ORGANISM: fusion
215 <400> SEQUENCE: 10
216   gactacaagg atgacgacga caaggccctg gataccaact actgcttcag ctccacggag      60
217   aagaactgct gctgcgga gctctacatt gacttcgga aggacctggg ctggaagtgg      120
218   attcatgaac ccaagggctc catgccaatt tctgcctggg gccctgtccc tacatctgga      180
219   gcctagacac tcagtacagc aaggtcctgg ctctgtacaa ccagcacaac ccgggcgcgt      240
220   cggcggcgcc gtgctgcgtg ccgcaggcgc tggagccact gcccatcggt tactacgtgg      300
221   gccgcaagcc caaggtggag cagctgtcca acatgatcgt gcgttcctgc aagtgcagct      360
222   ga
224 <210> SEQ ID NO: 11
225 <211> LENGTH: 120
226 <212> TYPE: PRT
227 <213> ORGANISM: fusionprotein
228 <220> FEATURE:
229 <221> NAME/KEY: PEPTIDE
230 <222> LOCATION: (1)..(8)
231 <223> OTHER INFORMATION: FLAG tag
232 <400> SEQUENCE: 11
233   Asp Tyr Lys Asp Asp Asp Asp Lys Ala Leu Asp Thr Asn Tyr Cys Phe
234   1               5               10               15
235   Ser Ser Thr Glu Lys Asn Cys Cys Val Arg Gln Leu Tyr Ile Asp Phe
236   20               25               30
237   Arg Lys Asp Leu Gly Trp Lys Trp Ile His Glu Pro Lys Gly Tyr His
238   35               40               45
239   Ala Asn Phe Cys Leu Gly Pro Cys Pro Tyr Ile Trp Ser Leu Asp Thr
240   50               55               60
241   Gln Tyr Ser Lys Val Leu Ala Leu Tyr Asn Gln His Asn Pro Gly Ala
242   65               70               75               80
243   Ser Ala Ala Pro Cys Cys Val Pro Gln Ala Leu Glu Pro Leu Pro Ile
244   85               90               95
245   Val Tyr Tyr Val Gly Arg Lys Pro Lys Val Glu Gln Leu Ser Asn Met
246   100              105              110
247   Ile Val Arg Ser Cys Lys Cys Ser
248   115              120
250 <210> SEQ ID NO: 12
251 <211> LENGTH: 1197

```

The type of errors shown exist throughout
the Sequence Listing. Please check subsequent
sequences for similar errors.

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 10/017,372A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length
The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering
The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length
Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES)
Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES)
Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000
- 9 Use of n's or Xaa's
 (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10 ✓ Invalid <213>
 Response
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220>
Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH - Biotechnology Systems Branch - 08/21/2001